Polarization imaging-based radiomics approach for the staging of liver fibrosis

YUE YAO,[1,2] FENGDI ZHANG,[1,2] BIN WANG,[3,4,5] JIACHEN WAN,[1,2] LU SI,[1,2] YANG DONG,[1,2] YUANHUAN ZHU,[1,2] XIAOLONG LIU,[6] LIHONG CHEN,[3,4,5,7] AND HUI MA[1,2,6,8]

[1]*Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Center for Precision Medicine and Healthcare, Shenzhen 518071, China*
[2]*Tsinghua Shenzhen International Graduate School, Shenzhen Key Laboratory for Minimal Invasive Medical Technologies, Institute of Optical Imaging and Sensing, Shenzhen 518055, China*
[3]*Fujian Medical University, Department of Pathology and Institute of Oncology, School of Basic Medical Sciences, Fuzhou 350014, China*
[4]*Fujian Medical University, Diagnostic Pathology Center, Fuzhou 350014, China*
[5]*Fujian Medical University, Mengchao Hepatobiliary Hospital, Fuzhou 350014, China*
[6]*Tsinghua University, Department of Physics, Beijing 100084, China*
[7]*drlhchen@sina.com*
[8]*mahui@tsinghua.edu.cn*

Abstract: Mueller matrix imaging contains abundant biological microstructure information and has shown promising potential in clinical applications. Compared with the ordinary unpolarized light microscopy that relies on the spatial resolution to reveal detailed histological features, Mueller matrix imaging encodes rich information on the microstructures even at low-resolution and wide-field conditions. Accurate staging of liver fibrosis is essential for the therapeutic diagnosis and prognosis of chronic liver diseases. In the clinic, pathologists commonly use semiquantitative numerical scoring systems to determine the stages of liver fibrosis based on the visualization of stained characteristic morphological changes, which require skilled staining technicians and well-trained pathologists. A polarization imaging based quantitative diagnostic method can help to reduce the time-consuming multiple staining processes and provide quantitative information to facilitate the accurate staging of liver fibrosis. In this study, we report a polarization imaging based radiomics approach to provide quantitative diagnostic features for the staging of liver fibrosis. Comparisons between polarization image features under a 4× objective lens with H&E image features under 4×, 10×, 20×, and 40× objective lenses were performed to highlight the superiority of the high dimensional polarization image features in the characterization of the histological microstructures of liver fibrosis tissues at low-resolution and wide-field conditions.

1. Introduction

Mueller matrix imaging contains abundant biological microstructure information [1–4] and has been applied in the characterization of various pathological tissues such as cervical precancerous lesions [5,6], breast cancerous tissues [7], and white matter fiber tracts of brain tissues [8]. The *in vivo* low-cost wide-field high definition Mueller polarimetric endoscope designed by Qi *et al*. [9,10] demonstrates the potential applications of Mueller matrix imaging in the wide-field early epithelial cancer diagnosis, surgical margin detection, and energy-based tissue fusion monitoring. Mueller matrix imaging is sensitive to biological microstructures even in sub-wavelength scales [1]. Compared with the commercial unpolarized light microscopy that relies on the spatial resolution to reveal detailed histological features, Mueller matrix imaging encodes sub-wavelength microstructures even at low-resolution and wide-field conditions [11,12].

In Mueller matrix imaging, previous studies have been focusing on using polarization parameters to characterize the pathological microstructures at pixel level rather than image level. For example, Wang *et al.* [13] used the average value of polarization parameters δ and t that involve minimal polarization image textures to characterize the liver fibrosis tissues. Dong *et al.* [7] proposed a pixel-based extraction approach to derive polarimetry feature parameters (PFPs) in the characterization of breast cancerous tissues. In recent years, several studies start to adopt the image texture analysis approach. Dong *et al.* [14] used the gray level co-occurrence matrix (GLCM) method in the quantitative characterization of breast ductal carcinoma tissues. Yao *et al.* [15] applied the local binary pattern (LBP) technique in the quantitative characterization of endometrial phases. Recently, Liu *et al.* [11] quantified the degree of correlation between H&E image texture features and pixel-based polarization parameter sets to provide additional microstructural information in assisting diagnosis. This research highlights the superiority of pixel-based polarization parameters in low-resolution and wide-field conditions. However, few previous studies involve systematic polarization parameter image based texture analysis.

Radiomics builds upon the advance in computer-aided clinical studies [16–19]. It transforms tumor pathophysiology information encrypted by medical images into high dimensional minable features to provide detailed quantification of tumor histological features [20,21]. Radiomics features include the intensity, shape, texture, and wavelet information of tumor images. Radiomics has been applied in various cancer studies to improve diagnostic, prognostic, and predictive accuracies. For example, Abdollahi *et al.* [22] applied radiomics approaches on magnetic resonance images to correlate the radiomics features with the outcomes of radiation therapy for the future prognosis of prostate cancers. In addition, in the era of precision medicine, radiomics features can be combined with other radiological, genomic, and proteomic information to establish objective, multi-dimensional and effective diagnostic and prognostic processes. Lambin *et al.* [21] summarized the workflow of radiomics into five phases: data selection, medical imaging, feature extraction, exploratory analysis, and modeling. Mueller matrix imaging, as an innovative quantitative imaging method that is high dimensional and contains rich microstructure information of pathological tissues, demonstrates the potential of developing a computer-aided polarization imaging based radiomics model in the quantitative characterization of pathological tissues.

Liver disease has become a major world public health problem [23,24]. Liver fibrosis is a characteristic of most types of chronic liver diseases. The accurate staging of liver fibrosis is essential for the therapeutic diagnosis and prognosis of chronic liver diseases. In the clinic, liver biopsy is the reference standard for the assessment of liver fibrosis [25]. Pathologists commonly use semiquantitative numerical scoring systems such as Scheuer [26], Metavir [27], Ishak [28], and Batt-Ludwig [29] to determine the stages of liver fibrosis based on the visualization of characteristic morphological changes of Hematoxylin and Eosin (H&E), Masson's trichrome, and Gordon and Sweet's silver stained sections. The multiple staining processes are time-consuming and require skilled technicians to obtain qualified staining sections for accurate diagnosis. The complex semiquantitative numerical scoring systems focus on qualitative histological features rather than quantitative histological features and require well-trained pathologists with rich experience in clinical diagnosis. Thus, a polarization imaging based quantitative diagnostic method can help to reduce the multiple staining processes and provide quantitative information to facilitate the accurate staging of liver fibrosis. Previous research [13] indicates the statistical analysis of Mueller matrix polar decomposition (MMPD) and Mueller matrix transformation (MMT) parameters can facilitate the preliminary quantification of 8-µm-thick non-stained and dewaxed sections of liver fibrosis tissues. Further study with more tissue samples is necessary to investigate the polarization imaging encoded texture features of liver fibrosis tissues. In this study, a polarization imaging based image feature analysis approach is presented to provide a

quantitative method that focuses on systematic polarization parameter image textures analysis in the staging of liver fibrosis.

2. Methods and materials

2.1. Experimental setup

In this study, a dual division of focal plane (DoFP) polarimeters-based full Mueller matrix microscope (DoFPs-MMM) as shown in Fig. 1 was used for fast full Mueller matrix imaging [30]. The DoFPs-MMM is designed by adding a polarization state generator (PSG) and a polarization state analyzer (PSA) on a commercial transmission microscope (L2050, Guangzhou LISS Optical Instrument Co., Ltd., China). The PSG contains a fixed-angle linear polarizer (P1) and a rotatable zero-order quarter-wave plate (R1). The four independent polarization states are generated by rotating the fast axis of R1 to four angles including $\pm 45°$ and $\pm 19.6°$. The PSA contains two 16 bit DoFP polarimeters (PHX050S-PC, Lucid Vision Labs Inc., Canada), a fixed-angle zero-order quarter-wave plate (R2), and a 50:50 non-polarized beam splitter prism (CCM1-BS013/M, Thorlabs Inc., USA). A DoFP polarimeter adds a pixelated micro-polarizer array (MPA) in front of an ordinary CCD sensor to form super pixels. Each super pixel contains four adjacent pixels with different polarization orientations, which enables the measurement of four linear polarization channels including 90°, 45°, 135°, and 0° in a single shot. The two DoFP polarimeters with the same field of view (FOV), resolution, and exposure time are fixed to the transmission and reflection ends of the 50:50 non-polarized beam splitter prism, and aligned with a polarization orientation of 0°. A fixed-angle zero-order quarter-wave plate (R2) is placed between the transmission end of the prism and DoFP-CCD1. During the experiment, the incident beam from the LED (633 nm, $\Delta\lambda = 20$ nm) is modulated by the PSG and then passes through the sample. The scattered light from the sample passes through the objective lens and is finally recorded by the two DoFP CCDs of PSA. The Mueller matrix of the sample ($\mathbf{M}_{sample}$) is calculated based on the equation as shown in Eq. (1) [30].

$$\mathbf{M}_{sample} = [\mathbf{S}_{out}][\mathbf{S}_{in}]^{-1} \tag{1}$$

where $[\mathbf{S}_{in}]$ contains four independent polarization states generated through the rotation of R1 in the PSG. And $[\mathbf{S}_{out}]$ refers to the polarization states detected by the PSA in a single shot. Detailed information relative to the DoFPs-MMM is demonstrated in [30].

2.2. Liver fibrosis tissue samples

This study involves 38 patients with pathologists determined stages of liver fibrosis. In this study, the samples were collected from patients who underwent surgeries for liver cancers between January 2020 to December 2021 in Mengchao Hepatobiliary Hospital. Among the 38 patients, there are 31 male patients and 7 female patients. The average age for male patients is 58.16 years old, and the average age for female patients is 61.43 years old. The inclusion of these 38 patients in this study mainly because they have liver biopsy tissues performed H&E, Masson's trichrome, and reticular fiber staining, and their stages of liver fibrosis tissues were determined by the Scheuer [31] grading system. This study excluded patients with incomplete clinical data, severe cardiovascular diseases, and other severe systematic diseases.

The Scheuer [31] system defines liver fibrosis into five stages. In stage 0 (S0), there is no fibrosis occurs. Stage 1 (S1) is defined with the appearance of the enlarged, fibrotic portal tracts. Stage 2 (S2) is defined with the formation of periportal or portal-portal septa in intact architecture. In stage 3 (S3), the fibrosis appears with architectural distortion, but no obvious cirrhosis. Stage 4 (S4) is considered probable or definite cirrhosis. In this study, among the 38 patients, a total of 10 patients, 10 patients, 9 patients, and 9 patients were diagnosed with stage 1 (S1), stage 2 (S2), stage 3 (S3), and stage (4) of liver fibrosis, respectively. For each patient, a
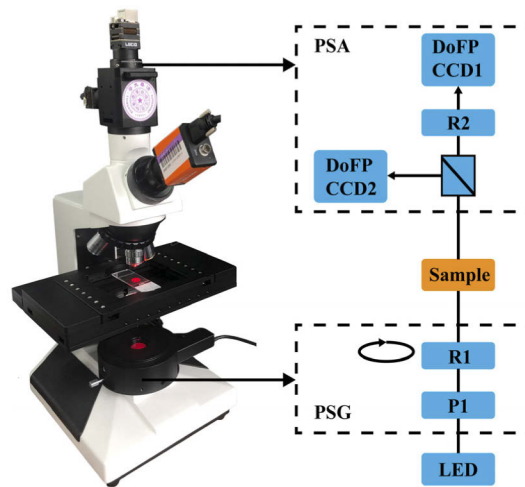
Fig. 1. Photograph and schematic of Mueller matrix microscope based on dual DoFP polarimeters (DoFPs-MMM).

4-µm-thick H&E-stained pathological slide was used to acquire polarization parameter images and H&E images. For each H&E-stained slide, one region of interest (ROI) that indicates the typical pathological features of liver fibrosis as shown in Fig. 2 was labeled by an experienced pathologist. Polarization imaging was performed on the labeled ROIs of the H&E-stained tissue sections to acquire polarization parameter images. And the corresponding H&E images under different resolutions were also obtained for comparisons. Figure 2 shows the ROIs of the H&E images for liver fibrosis tissues from S1 to S4 under objective lenses of 4×, 10×, 20×, and 40×. This study strictly followed the rules regulated by the Ethics Committee of the Mengchao Hepatobiliary Hospital.

2.3. Polarimetry basis parameters (PBPs)

Mueller matrix contains full polarization information of pathological samples, but the individual elements often lack explicit connections to specific microstructures. Therefore, various polarization parameters [32] with more explicit physical meanings were derived from the Mueller matrix to characterize the microstructures of pathological samples. In recent research, Dong *et al*. [7] summarized previously derived polarization parameters from different physical methods into Polarimetry Basis Parameters (PBPs). This study also employed the term Polarimetry Basis Parameter (PBP) to characterize the microstructures of liver fibrosis tissues. Table S1 of the supplemental document summarizes the calculation equations and physical meanings of the 26 PBPs used in this study. The 26 PBPs including 5 Mueller matrix polar decomposition (MMPD) parameters [32–34], 5 Mueller matrix transformation (MMT) parameters [1,32], 4 Mueller matrix rotation invariant (MMRI) parameters along with 4 corresponding orientation angle parameters [32,35], 4 Mueller matrix linear birefringence identity (MMLBI) parameters [5,32] and 4 Mueller matrix linear diattenuation identity (MMLDI) parameters [5,32]. The 5 MMPD parameters are linear retardance parameter δ, linear retardance orientation angle parameter θ, diattenuation parameter D, depolarization parameter Δ, and the optical rotation parameter Ψ. The 5 MMT parameters include anisotropy degree parameter t_1, polarizance parameter b, circular birefringence parameter β, normalized anisotropy parameter A and local orientation parameter α. The 4 MMRI parameters contain linear polarizance parameter P_L, linear diattenuation parameter D_L, and the birefringence related parameters q_L and r_L that also describe the capabilities of transforming
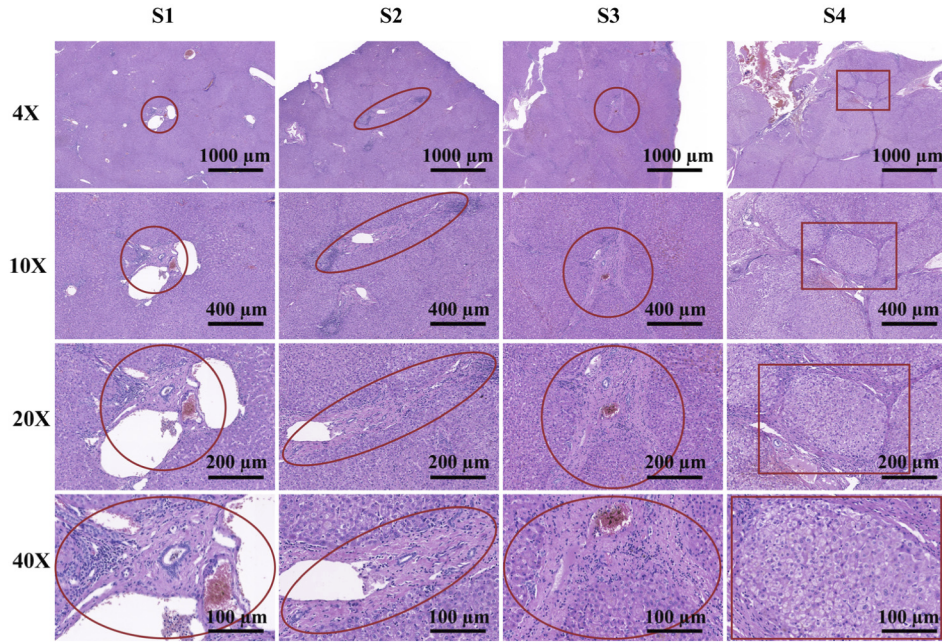
Fig. 2. The 4-µm-thick H&E-stained pathological features of liver fibrosis tissues at stages S1 to S4 under 4×, 10×, 20×, and 40× objective lenses. The typical liver fibrosis histological features for each stage were highlighted in red marks.

between linear and circular polarizations. The 4 corresponding orientation angle parameters involve linear polarizance orientation angle parameter α_P, linear diattenuation orientation angle parameter α_D, and birefringence related orientation angle parameters α_q and α_r. The 4 MMLBI parameters are P_1, P_2, P_3 and P_4 that derived through comparison between an asymmetric anisotropic Mueller matrix with a pure linear retarder Mueller matrix. The 4 MMLDI parameters are P_5, P_6, P_7 and P_8 that derived through comparison between an asymmetric anisotropic Mueller matrix with a pure linear diattenuator Mueller matrix.

2.4. Workflow of polarization parameter images based radiomics approach

The workflow of polarization imaging based image feature analysis approach includes four steps: (1) polarization parameter images acquisition; (2) polarization parameter images feature extraction; (3) polarization parameter images feature selection and classification; (4) and the interpretation and correlation of selected significant polarization parameter image features with the microstructures of liver fibrosis tissues from S1 to S4. Figure 3 illustrates the workflow of the polarization imaging based radiomics approach in detail.

In the polarization parameter image feature extraction step as shown in the workflow of Fig. 3, for a single ROI of a liver fibrosis sample, the inputs were 26 Polarization Basis Parameter (PBP) images derived from different physical methods with explicit physical meanings. For each PBP image, a total of 110 image features were extracted. Therefore, for a single ROI of a liver fibrosis sample, as many as 2860 PBP image features can be extracted. The 26 PBPs were introduced in Table S1 of the supplemental document in detail. The 110 image features [36] include: (1) 18 intensity based statistical features; (2) 23 discretised intensity frequency count histogram (DIFCH) based features; (3) 6 intensity volume histogram (IVH) based features; (4) 25 gray level co-occurrence matrix (GLCM) based features; (5) 16 gray level run length matrix (GLRLM) based features; (6) 5 neighborhood gray tone difference matrix (NGTDM) based features; (7) and
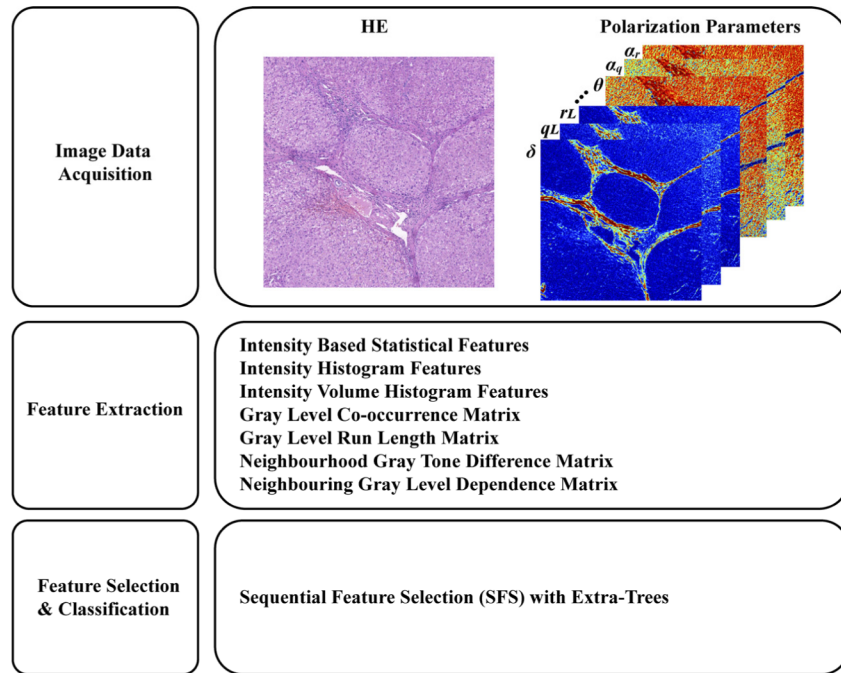
Fig. 3. Workflow of polarization parameter images based radiomics approach for the staging of liver fibrosis.

17 neighboring gray level dependence matrix (NGLDM) based features. The detailed names, meanings, and calculation equations of the 110 image features were shown in the reference manual published by the image biomarker standardization initiative (IBSI) [36].

In the polarization parameter image feature selection and classification step as shown in the workflow of Fig. 3, the forward Sequential Feature Selection (SFS) [37] was combined with the Extra-Trees classifier to assess the performance of the 26 Polarization Basis Parameters at the image features level. In the combination of forward SFS with Extra-Trees, the forward SFS method sequentially adds the best polarization parameter image feature based on the highest 5-fold cross-validation scores of the Extra-Trees classification on liver fibrosis tissues at four stages to form a selected significant features subset in a greedy fashion [38]. The Extra-Trees classifier imported from the "sklearn.ensemble.ExtraTreesClassifier" of the Scikit-Learn library [38] is an Extremely Randomized Trees ensemble of Decision Trees, which are suitable for fitting complex datasets in classification, regression, and multi-output tasks. During the machine learning classification, the Decision Trees create sequential questions that split the datasets into smaller groups and a predictive decision is made at the terminal node. In Decision Trees, Gini impurity that measures the probability of misclassification is used to decide the optimal split from a root node or subsequent splits, and a lower Gini impurity represents a better split [39]. In this study, the combination of forward Sequential Feature Selection with Extra-Trees classifier simultaneously outputs the highest liver fibrosis four-class classification accuracy with its corresponding subset of significant and interpretable PBP image features that play a critical role in the characterization of the microstructural features of liver fibrosis tissues.

3. Results and discussion

3.1. Polarization imaging results of liver fibrosis tissue samples

Previous studies indicate polarization imaging is sensitive to pathological changes of fibrous structures [13,14]. Dong *et al.* [14] state the MMPD linear retardance parameter δ and its corresponding orientation angle parameter θ are sensitive to the fibrous structures of breast ductal carcinoma tissues. Wang *et al.* [13] indicate that birefringence related polarization parameters play an important role in the characterization of liver fibrosis tissues. He *et al.* [1] summarize the birefringence related polarization parameters δ, q_L and r_L are more sensitive to collagen and fibrous structures. The diattenuation related polarization parameters P_L and D_L are more sensitive to cell nucleus structures. In this study, the polarization imaging based radiomics approach selected 8 significant polarization parameters including the linear birefringence related parameter P_1 and P_4, the diattenuation related parameter P_5 and D, the polarizance parameter b, the circular birefringence parameter β, the linear diattenuation orientation related parameter α_P, and the linear birefringence orientation related parameter α_r. In Fig. 4, example images of selected significant parameters P_4, D, α_P and α_r of liver fibrosis tissues at stages S1, S2, S3, and S4 under a 4× objective lens were shown.
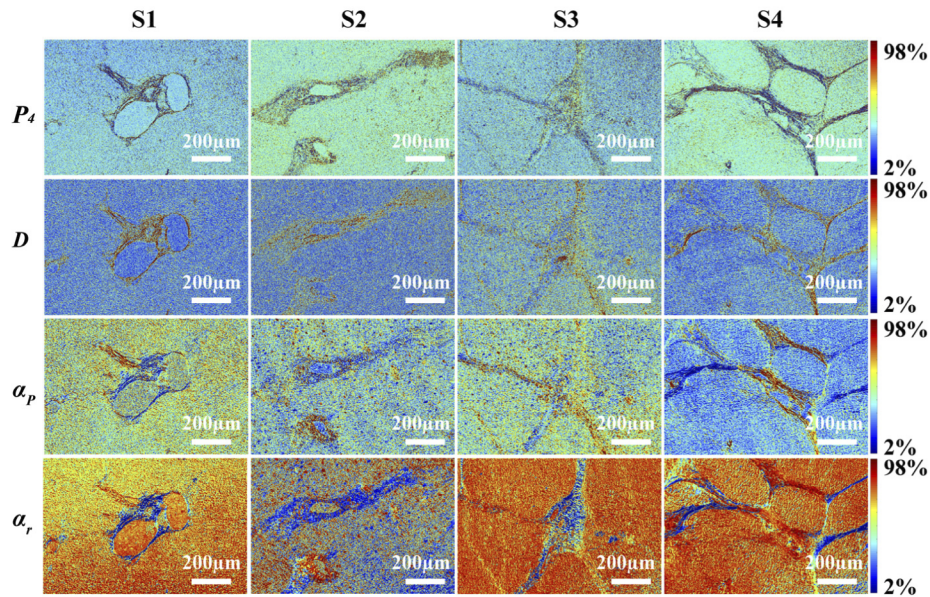

Fig. 4. Images of Polarimetry Basis Parameters P_4, D, α_P and α_r of liver fibrosis tissues at stages S1, S2, S3 and S4. The image size is 1591× 2411 pixels. The color bar of the image ranges from the 2nd percentile to the 98th percentile of the polarization parameter.

3.2. Polarization image features selection with forward sequential feature selection

The combination of forward Sequential Feature Selection (SFS) with Extra-Trees classification [37] was applied to the 2860 PBP image features to acquire the best PBP image features subset with the highest classification accuracy in the staging of liver fibrosis tissues. This combination sequentially selects the best polarization parameter image feature based on the highest 5-fold cross-validation scores of Extra-Trees classifier to form a selected features subset in a greedy fashion [38]. During the machine learning based PBP image feature selection and classification process, the 5-fold cross-validation was added for the selection of PBP image features and the

classification of liver fibrosis tissues at four stages. This combined process simultaneously outputs the selected PBP image features subset along with their corresponding four-class liver fibrosis classification accuracy, which was calculated by taking the mean of the 5-fold cross-validation accuracies. To add the 5-fold cross-validation, the 38 liver fibrosis samples were randomly separated into 5 folds with a relatively even distribution of classes. During the 5-fold cross-validation, 4 folds were used as the training set and 1 fold was used as the testing set. The cross-validation repeats until each fold of the 5 folds has been used as the testing set. The four-class liver fibrosis classification accuracy in this study was obtained by taking the mean of the 5 testing set accuracies of the 5-fold cross-validation. Figure 5 shows the four-class liver fibrosis classification accuracy from 1 PBP image feature up to 15 PBP image features with an increment of 1. The results indicate a subset of 12 PBP image features achieved the highest four-class classification accuracy of 94.6%.
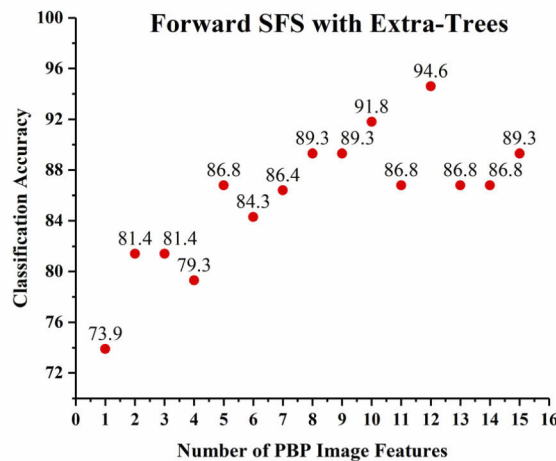

Fig. 5. Four-class liver fibrosis classification accuracies of selected PBP image features from 1 PBP image feature up to 15 PBP image features with an increment of 1.

The confusion matrices of the 12 PBP image features subset with the highest accuracy were shown in Fig. 6. Figure 6(a) to Fig. 6(e) show the confusion matrix of using each fold of the 5 folds as the testing set in the 5-fold cross-validation process. The results indicate the testing set accuracy for fold 1, fold 2 and fold 4 was all 100% as shown in Fig. 6(a), Fig. 6(b) and Fig. 6(d), respectively. The testing set accuracy for fold 3 and fold 5 was 87.50% and 85.71% as shown in Fig. 6(c) and Fig. 6(e) respectively. For fold 3, one liver fibrosis sample at S3 was misclassified into S1. For fold 5, one liver fibrosis sample at S4 was misclassified into S3.

To sum up, among the total of 38 liver fibrosis samples, the 10 samples from S1 and the 10 samples from S2 were all classified correctly. For samples from S3 and S4, both 8 out of 9 samples were classified correctly. The one sample in S3 misclassified into S1 causes greater consequences than the one sample in S4 misclassified into S3 in terms of affecting the later clinical treatment and prognosis. One of the causes of the misclassification was the limitation of sample sizes. In the future, more liver fibrosis samples will be collected to improve the classification accuracies.

Table 1 demonstrates the details of the selected PBP image features from 1 feature up to 15 features in increasing order with corresponding four-class liver fibrosis classification accuracies shown in Fig. 5. The first 12 PBP image features were significant PBP image features that can form a subset to achieve the highest four-class liver fibrosis classification accuracy of 94.6% with the 5-fold cross-validation. In the polarization perspective, for the 12 PBP image features subset,
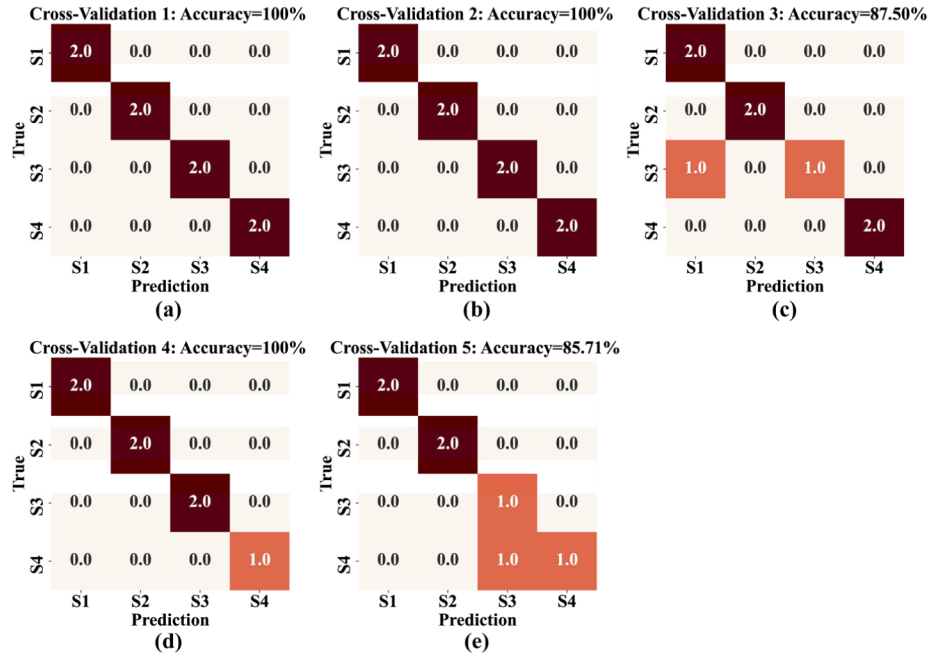
Fig. 6. The confusion matrix for each fold of the 5-fold cross-validation of the 12 PBP image features subset with the highest four-class classification accuracy of 94.6%. Figure 6(a) to Fig. 6(e) show the confusion matrix of using fold 1, fold 2, fold 3, fold 4 and fold 5 as the testing set, respectively.

a total of 8 significant PBPs including α_P, P_1, P_5, α_r, b, β, P_4, and D were selected from the 26 PBPs. The first PBP selected by the combined machine learning process was the orientation angle parameter α_P, which describes the orientations of diattenuation. The Mueller matrix linear birefringence identity (MMLBI) parameters P_1 and P_4 are the deviations of an asymmetric anisotropic Mueller matrix compared with a pure linear retarder Mueller matrix. Parameter P_1 measures the breaking of the symmetry of the Mueller matrix elements m_{43} and m_{34}. Parameter P_4 is obtained through subtracting the multiply of m_{42} and m_{43} from the multiply of m_{34} and m_{24}. The selection of parameters P_1 and P_4 indicate the appearance of liver fibrosis microstructures that break the symmetries of birefringence related Mueller matrix elements. The Mueller matrix linear diattenuation identity (MMLDI) parameter P_5 is a deviation of an asymmetric Mueller matrix compared with a pure linear diattenuator Mueller matrix. The selection of P_5 indicates the appearance of liver fibrosis microstructures that break the symmetry of linear diattenuation related Mueller matrix elements m_{12} and m_{21}. The orientation angle parameter α_r is related to the orientations of linear birefringence and is calculated by Mueller matrix elements m_{24} and m_{34}. The Mueller matrix transformation (MMT) parameter b is the polarizance related parameter calculated from Mueller matrix elements m_{22} and m_{33}. Parameter β is the circular birefringence parameter calculated from Muller matrix elements m_{23} and m_{32}. The Mueller matrix polar decomposition (MMPD) parameter D describes the diattenuation of liver fibrosis tissues. To summarize in the polarization perspective, the values of linear birefringence (P_1, P_4), diattenuation (P_5, D), polarizance (b), and circular birefringence (β) related parameters played important roles in distinguishing the liver fibrosis histological structures with the highest classification accuracy of 94.6%. The orientation angle parameters related to linear diattenuation (α_P) and linear birefringence (α_r) also made contributions in characterizing the liver fibrosis tissues at four stages.

Table 1. Selected PBP Image Features Subsets by the Combination of Forward SFS with Extra-Trees

Number	PBP	Image Features	
		List of 15 Selected PBP Image Features	
		Image Feature Name	Image Feature Meaning [36]
1	α_P	NGLDM Low Gray Level Count Emphasis (LGLCE)	Distribution of low gray-level values
2	P_1	DIFCH Mean Discretised Intensity	Mean of the discretised intensity
3	P_5	GLCM Difference Entropy	Randomness in neighborhood intensity value differences
4	α_P	GLRLM Low Gray Level Run Emphasis (LGLRE)	Distribution of low gray-level run lengths
5	α_r	NGTDM Complexity	An image is complex when it is non-uniform and has many rapid changes in gray-level intensity
6	b	GLCM Correlation	Joint probability occurrence of the specified pixel pairs
7	α_P	GLRLM Short Run Low Gray Level Emphasis (SRLGLE)	Joint distribution of short run lengths with low gray-level values
8	β	GLRLM Run Percentage	Coarseness of the texture calculated by taking the ratio of number of runs and number of pixels in the ROI
9	P_4	Intensity Based Energy	Measures the magnitude of pixel values in an image with a sum of the squares function
10	P_5	GLRLM Long Run Low Gray Level Emphasis (LRLGLE)	Joint distribution of long run lengths with low gray-level values
11	b	GLCM Information Correlation 2	Complexity of the texture
12	D	GLCM Correlation	Joint probability occurrence of the specified pixel pairs
13	α	DIFCH Discretised Intensity Kurtosis	Peakedness of the discretised intensity
14	α	NGLDM Low Dependence Low Gray Level Emphasis (LDLGLE)	Joint distribution of low dependence with low gray-level values
15	q_L	GLRLM Normalised Run Length Non-Uniformity (NRLNU)	Homogeneity of run lengths throughout the image

Previous studies on pathological tissues such as breast ductal carcinoma tissues [14] and liver fibrosis tissues [13] indicate polarization imaging is sensitive to fibrous structures. In the liver fibrosis study, Wang *et al.* [13] applied Mueller matrix imaging on four slices of 8-µm-thick unstained and dewaxed liver fibrosis tissues from stage F1 to stage F4 for quantitative microstructural characterization. The results indicate the MMPD linear retardance parameter δ and its corresponding orientation angle parameter θ, the retardance related MMT parameter t and its corresponding orientation angle parameter x are all can be used for the quantification of liver fibrosis tissues from F1 to F4. Wang *et al.* [13] also performed Monte Carlo simulations based on the sphere birefringence model to study the influence of the liver fibrosis progress on polarization parameters. The simulations demonstrate the progression of liver fibrosis causes the increase of birefringence values. Both experimental results and Monte Carlo simulations indicate that birefringence plays an important role to characterize liver fibrosis at four stages. Therefore, in this study, the selection of significant PBPs that are related to the values (P_1, P_4) and orientations (α_r) of birefringence in the classification of liver fibrosis tissues match with previous studies in terms of physical meanings. For thin pathological tissue sections, He *et al.* [1] summarize the birefringence related polarization parameters δ, q_L and r_L are more sensitive to collagen and fibrous structures. The diattenuation related polarization parameters P_L and D_L

are more sensitive to cell nucleus structures. Therefore, the selection of significant PBPs related to values (P_5, D) and orientations (α_P) of diattenuation may indicate the contributions of cell nucleus structural differences in the classification of liver fibrosis tissues. In addition, this study shows the image features of the polarizance parameter b and the circular birefringence parameter β also can make contributions to distinguish liver fibrosis tissues.

In the polarization image features perspective, as shown in Table 1, for the 12 PBP image features subset with the highest classification accuracy, parameter α_P that related to the orientations of linear diattenuation selected three image textures including the NGLDM based Low Gray Level Count Emphasis (LGLCE), the GLRLM based Low Gray Level Run Emphasis (LGLRE), and the GLRLM based Short Run Low Gray Level Emphasis (SRLGLE). Table 1 shows all of these three image textures measure the distribution of low gray-level values in parameter α_P, which indicates the distribution of low gray-level values in parameter α_P is particularly informative in characterizing the liver fibrosis microstructures from S1 to S4. The Mueller matrix linear birefringence identity parameters P_1 and P_4 selected the Intensity Histogram based Mean and the Intensity based Energy respectively. The Intensity based Energy measures the magnitude of pixel values in parameter P_4 with a sum of the squares function. The Mueller matrix linear diattenuation identity parameter P_5 selected two image textures including the GLCM based Difference Entropy and the GLRLM based Long Run Low Gray Level Emphasis (LRLGLE). The GLCM based Difference Entropy quantifies the randomness in neighborhood intensity value differences of parameter P_5. The GLRLM based LRLGLE measures the joint distribution of long run lengths with low gray-level values in P_5. The linear birefringence orientation related angle parameter α_r selected the NGTDM based Complexity that measures the nonuniformity and rapid changes in gray-level intensities. The polarizance parameter b selected two image textures from the GLCM family including the Correlation and the Information Correlation 2. The Correlation describes the joint probability occurrence of the specified pixel pairs and the Information Correlation 2 measures the Complexity of the textures of parameter b. The circular birefringence related parameter β selected the GLRLM based Run Percentage that quantifies the Coarseness of the textures. The MMPD diattenuation parameter D also selected the GLCM based Correlation to distinguish liver fibrosis tissues from S1 to S4.

In summary, all the 12 significant PBP image features listed in Table 1 made contributions to achieving the highest four-class liver fibrosis classification accuracy of 94.6%. For the birefringence related parameters P_1 and P_4, the intensity histogram based Mean and the magnitude of intensities were useful. For the birefringence orientation related parameter α_r, the complexity and non-uniformity of the images were informative. For the diattenuation related parameter P_5, the randomness in neighborhood intensity value differences and the joint distribution of long run lengths with low gray-level values were important. For diattenuation related parameter D, the GLCM based Correlation was useful. To be noted, for the diattenuation orientation related parameter α_P, the distribution of low gray-level values was particularly informative because α_P selected 3 image features that describe the distribution of low gray-level values. Besides, the GLCM based correlation and complexity of polarizance parameter b, the coarseness of circular birefringence parameter β were all provided useful information for the quantitative characterization of liver fibrosis tissues at four stages. In addition, this study has a limited sample size of 38 patients. Even though with the addition of the 5-fold cross-validation process, the classification accuracy was relatively stable and robust and the selected significant PBPs were relatively consistent with our previous studies in terms of physical meanings. Further studies with more liver fibrosis samples also can be done in the future to improve the classification accuracies for clinical applications and further validate the polarization based radiomics approach.

3.3.   Comparison of significant PBP image features and H&E image features

Several preliminary studies have been investigated the effects of resolutions on polarization parameter images and H&E images. Shen *et al*. [12] studied the effects of resolution on MMPD derived linear retardance parameter δ and its corresponding orientation angle parameter θ. This study obtained the parameters δ and θ under 4× and 40× objective lenses. Then, six polarization parameter image features under 4× and 40× resolutions were calculated to perform a quantitative comparison. The results indicate the breast tissue fiber density information contained by the GLCM features of parameter δ was well preserved when the polarization imaging resolution decreases. But the breast tissue fiber orientation information contained by the orientation angle parameter θ was not completely preserved when the polarization image resolution decreases. Liu *et al*. [11] performed a comparison between pixel-based polarization parameters and H&E image textures in the characterization of breast ductal carcinoma tissues under multi-resolution cases. The results indicate for Tamura related H&E image textures, the accuracies to identify the target structures decreased significantly with the decline of resolutions. However, for polarization parameter Δ, t_1, D_L, δ and r_L, the accuracies to identify the target structures were stable with only slight fluctuations under multi-resolution cases. These preliminary studies demonstrate the performance of H&E image textures decreases significantly with the decline of resolutions. In comparison with H&E image textures, the performance of polarization parameters is less affected by the decline of resolutions. Therefore, in this study, comparisons were made between 4× low-resolution polarization parameter image features and multi-resolution H&E image features to demonstrate the advantage of polarization imaging under low-resolution and wide-field conditions in the staging of liver fibrosis.

In the comparative study, the radiomics approach was applied on H&E images under 4×, 10×, 20×, and 40× resolutions as shown in Fig. 2. The workflow of radiomics approach for H&E images includes (1) multi-resolution H&E images acquisition; (2) extraction of 110 image features for each channel (RGB) of H&E images; (3) and H&E image features selection and classification through the combination of forward SFS with Extra-Trees. This workflow used the same machine learning based feature selection, classification, and 5-fold cross-validation methods compared with the workflow of polarization parameter images that obtained the highest classification accuracy of 94.6%. The highest classification accuracies of PBP image features under a 4× resolution and H&E image features under 4×, 10×, 20×, and 40× resolutions were shown in Table 2.

Table 2. Classification Accuracy of 4 × PBP Image Features and Multi-Resolution H&E Image Features

	H&E 4 ×	H&E 10 ×	H&E 20 ×	H&E 40 ×	PBP 4 ×
Accuracy	63.9	71.4	73.6	77.1	94.6

The highest classification accuracy of PBP image features under a 4× resolution was 94.6%. The highest classification accuracies of H&E image features under 4×, 10×, 20×, and 40× resolutions were 63.9%, 71.4%, 73.6%, and 77.1% respectively. For H&E image features, the classification accuracy decreases with the decline of resolutions. From 40× to 20×, the accuracy decreased by 3.5%. From 20× to 10×, the accuracy decreased by 2.2%. And from 10× to 4×, the accuracy decreased significantly by 7.5%. In comparison with H&E image features under multi-resolutions, the significant PBP image features subset under a 4× low-resolution can achieve the highest classification accuracy of 94.6% in the characterization of liver fibrosis tissues from S1 to S4. The H&E images only contain 3 channels (Red channel, Green channel, and Blue channel). The polarization parameter images are high dimensional and contain 26 PBPs with more explicit physical meanings. These results demonstrate the high dimensional

polarization parameter images contain rich microstructural information of liver fibrosis tissues even at low-resolution and wide-field conditions.

3.4. Interpretation of significant H&E image features

The significant PBP image features under a 4× resolution were already discussed in detail in Table 1. The significant multi-resolution H&E image features corresponding to the classification accuracies shown in Table 2 were listed in detail in Table 3, Table 4, Table 5, and Table 6, respectively.

Table 3. Interpretation of Significant 4 × H&E Image Features Subset

Number	RGB Channel	4 × H&E Image Feature Name	4 × H&E Image Feature Meaning [36]
1	Blue	DIFCH Discretised Intensity Skewness	Asymmetry of discretised intensity
2	Red	DIFCH Intensity Histogram Median Absolute Deviation	Dispersion from the Median
3	Red	GLRLM Gray Level Variance (GLV)	Variance in gray level intensity for the runs

Table 4. Interpretation of Significant 10 × H&E Image Features Subset

Number	RGB Channel	10 × H&E Image Feature Name	10 × H&E Image Feature Meaning [36]
1	Blue	Skewness	Asymmetry of intensity
2	Red	DIFCH Median discretised intensity	Median of discretised intensity
3	Red	GLRLM Gray Level Variance (GLV)	Variance in gray level intensity for the runs

Table 5. Interpretation of Significant 20 × H&E Image Features Subset

Number	RGB Channel	20 × H&E Image Feature Name	20 × H&E Image Feature Meaning [36]
1	Red	NGTDM Strength	A measure of primitives, value is high when the primitives are easily defined and visible
2	Red	GLRLM Run Length Variance (RLV)	Variance in runs for the run lengths
3	Red	Mean	Mean of intensity
4	Red	GLCM Inverse Difference Moment	Local homogeneity
5	Red	GLRLM Short Run Emphasis (SRE)	Distribution of short run lengths, value is high for fine textures
6	Red	DIFCH P10	10th discretised intensity percentile

For the H&E images under a 4× resolution, three image features as shown in Table 3 were selected to achieve the highest four-class classification accuracy of 63.9%. The results indicate for 4× H&E images, the intensity histogram (DIFCH) based features that measure the asymmetry of the blue channel and the dispersion from the median of the red channel contain informative textural information to distinguish liver fibrosis tissues at four stages. The GLRLM based Gray Level Variance that measures the variance in gray level intensity for the runs of the red channel also provided additional information for the characterization.

For the H&E images under a 10× resolution, three image features as shown in Table 4 were selected to achieve the highest four-class classification accuracy of 71.4%. For 10× H&E images, the intensity based statistical feature of Skewness that measures the intensity asymmetry of the blue channel, the intensity histogram (DIFCH) based feature that measures the Median of

Table 6. Interpretation of Significant 40 × H&E Image Features Subset

Number	RGB Channel	40 × H&E Image Feature Name	40 × H&E Image Feature Meaning [36]
1	Blue	GLCM Difference Entropy	Randomness in neighborhood intensity value differences
2	Red	DIFCH Minimum Histogram Gradient	Minimum Histogram Gradient
3	Red	GLCM Difference Entropy	Randomness in neighborhood intensity value differences
4	Blue	DIFCH Minimum Histogram Gradient Intensity	Discretised intensity corresponding to the Minimum Histogram Gradient
5	Green	DIFCH P10	10th discretised intensity percentile
6	Green	DIFCH Intensity Histogram Median Absolute Deviation	Dispersion from the Median
7	Red	GLRLM Short Run High Gray Level Emphasis (SRHGLE)	Joint distribution of short run lengths with high gray-level values
8	Blue	GLRLM Long Run Emphasis (LRE)	Distribution of long run lengths, value is high for coarse structural textures

discretised intensity of the red channel, and the GLRLM based GLV that measures the variance in gray level intensity for the runs are all informative to classify the liver fibrosis tissues.

For the H&E images under a 20× resolution, six image features as shown in Table 5 were selected to achieve the highest four-class classification accuracy of 73.6%. The six image features were all extracted from the red channel of the 20× H&E images. The informative image features to distinguish the liver fibrosis tissues include the NGTDM based Strength that measures the image primitives, the GLRLM based RLV that quantifies the variance in runs for the run lengths, the intensity based statistical feature of Mean, the GLCM based Inverse Difference Moment that measures the local homogeneity, the GLRLM based SRE that describe the fineness of image textures, and the intensity histogram (DIFCH) based feature of 10th discretised intensity percentile.

For the H&E images under a 40× resolution, 8 image features as shown in Table 6 were selected to achieve the highest four-class classification accuracy of 77.1%. Three image features extracted from the blue channel provided useful textural information to classify the liver fibrosis tissues. They were the GLCM based Difference Entropy that quantifies the randomness in neighborhood intensity value differences, the intensity histogram (DIFCH) based Minimum Histogram Gradient Intensity, and the GLRLM based Long Run Emphasis (LRE) that measures the coarseness of image textures. Besides, three image features extracted from the red channel were also selected for classification. They were the intensity histogram (DIFCH) based Minimum Histogram Gradient, the GLCM based Difference Entropy that related to the randomness of the red channel, and the GLRLM based Short Run High Gray Level Emphasis (SRHGLE) that describes the joint distribution of short run lengths with high gray-level values. The two green channel extracted intensity histogram (DIFCH) based features that calculate the 10th discretised intensity percentile and the dispersion from the Median also provided important textural information to achieve a four-class classification accuracy of 77.1%.

The comparisons of the 12 significant PBP image features listed in Table 1 with the multi-resolution significant H&E image features listed in Table 3, Table 4, Table 5, and Table 6 do not show explicit correlations and tendencies in the characterization of the histological structures of liver fibrosis tissues. Only one image feature was selected both by the 4× polarization parameter images and the multi-resolution H&E images, which was the GLCM based Difference Entropy. It measures the randomness in neighborhood intensity value differences and was selected by both

the diattenuation related parameter P_5 under a 4× resolution and the H&E images under a 40× resolution.

4. Conclusions

In conclusion, we systematically analyzed the image features of 26 polarization parameters with explicit physical meanings through the radiomics approach. This study demonstrates the feasibility to provide a polarization imaging based radiomics approach that focuses on the quantitative image features analysis in the clinical staging of liver fibrosis. In addition, comparisons between polarization image features under a 4× objective lens with H&E image features under 4×, 10×, 20×, and 40× objective lenses were performed. With the same radiomics approach, the highest classification accuracies of H&E image features under 4×, 10×, 20×, and 40× resolutions were 63.9%, 71.4%, 73.6%, and 77.1% respectively. The highest classification accuracy of polarization parameter images under a 4× resolution was 94.6%. In this study, 8 significant polarization parameters were selected out of 26 polarization parameters to achieve the highest classification accuracy of 94.6%. The results indicate the 8 PBP image features extracted from the linear birefringence (P_1, P_4), diattenuation (P_5, D), polarizance (b), and circular birefringence (β) related parameters played important roles in distinguishing the liver fibrosis histological structures from S1 to S4. The 4 PBP image features extracted from the orientation angle parameters related to linear diattenuation (α_P) and linear birefringence (α_r) also made contributions in characterizing the liver fibrosis tissues. The comparison of the significant PBP image features with the significant H&E image features does not show explicit correlations and tendencies in the staging of liver fibrosis tissues. Only one image feature was selected by both the 4× polarization parameter images and the multi-resolution H&E images, which was the GLCM based Difference Entropy that measures the randomness in neighborhood intensity value differences. It was selected by both the diattenuation related parameter P_5 under a 4× resolution and the H&E images under a 40× resolution. In conclusion, this study demonstrates that high dimensional polarization parameter image features are good at quantitatively characterizing the microstructural features of liver fibrosis tissues even at low-resolution and wide-field conditions. The superiority of characterizing the histological structures in low-resolution and wide-field conditions enables the potential to integrate polarization imaging in various *in vivo* and endoscopic imaging systems to facilitate the clinical diagnosis and prognosis.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Supplemental document. See Supplement 1 for supporting content.

References

1. H. He, R. Liao, N. Zeng, P. Li, Z. Chen, X. Liu, and H. Ma, "Mueller matrix polarimetry—an emerging new tool for characterizing the microstructural feature of complex biological specimen," J. Lightwave Technol. **37**(11), 2534–2548 (2019).
2. C. He, H. He, J. Chang, B. Chen, H. Ma, and M. J. Booth, "Polarisation optics for biomedical and clinical applications: a review," Light: Sci. Appl. **10**(1), 1–20 (2021).
3. S. L. Jacques, J. C. Ramella-Roman, and K. Lee, "Imaging skin pathology with polarized light," J. Biomed. Opt. **7**(3), 329–340 (2002).
4. S. Alali and I. A. Vitkin, "Polarized light imaging in biomedicine: emerging Mueller matrix methodologies for bulk tissue assessment," J. Biomed. Opt. **20**(6), 061104 (2015).
5. Y. Dong, J. Wan, X. Wang, J.-H. Xue, J. Zou, H. He, P. Li, A. Hou, and H. Ma, "A polarization-imaging-based machine learning framework for quantitative pathological diagnosis of cervical precancerous lesions," IEEE Trans. Med. Imaging **40**(12), 3728–3738 (2021).

6. M. Sun, H. He, N. Zeng, E. Du, Y. Guo, S. Liu, J. Wu, Y. He, and H. Ma, "Characterizing the microstructures of biological tissues using Mueller matrix and transformed polarization parameters," Biomed. Opt. Express **5**(12), 4223–4234 (2014).

7. Y. Dong, J. Wan, L. Si, Y. Meng, Y. Dong, S. Liu, H. He, and H. Ma, "Deriving polarimetry feature parameters to characterize microstructural features in histological sections of breast tissues," IEEE Trans. Biomed. Eng. **68**(3), 881–892 (2021).

8. P. Schucht, H. R. Lee, H. M. Mezouar, E. Hewer, A. Raabe, M. Murek, I. Zubak, J. Goldberg, E. Kövari, and A. Pierangelo, "Visualization of white matter fiber tracts of brain tissue sections with wide-field imaging Mueller polarimetry," IEEE Trans. Med. Imaging **39**(12), 4376–4382 (2020).

9. J. Qi and D. S. Elson, "A high definition Mueller polarimetric endoscope for tissue characterisation," Sci. Rep. **6**(1), 1–11 (2016).

10. J. Qi and D. S. Elson, "Mueller polarimetric imaging for surgical and diagnostic applications: a review," J. Biophotonics **10**(8), 950–982 (2017).

11. Y. Liu, Y. Dong, L. Si, R. Meng, Y. Dong, and H. Ma, "Comparison between image texture and polarization features in histopathology," Biomed. Opt. Express **12**(3), 1593–1608 (2021).

12. Y. Shen, R. Huang, H. He, S. Liu, Y. Dong, J. Wu, and H. Ma, "Comparative study of the influence of imaging resolution on linear retardance parameters derived from the Mueller matrix," Biomed. Opt. Express **12**(1), 211–225 (2021).

13. Y. Wang, H. He, J. Chang, C. He, S. Liu, M. Li, N. Zeng, J. Wu, and H. Ma, "Mueller matrix microscope: a quantitative tool to facilitate detections and fibrosis scorings of liver cirrhosis and cancer tissues," J. Biomed. Opt. **21**(7), 071112 (2016).

14. Y. Dong, J. Qi, H. He, C. He, S. Liu, J. Wu, D. S. Elson, and H. Ma, "Quantitatively characterizing the microstructural features of breast ductal carcinoma tissues in different progression stages by Mueller matrix microscope," Biomed. Opt. Express **8**(8), 3643–3655 (2017).

15. Y. Yao, M. Zuo, Y. Dong, L. Shi, Y. Zhu, L. Si, X. Ye, and H. Ma, "Polarization imaging feature characterization of different endometrium phases by machine learning," OSA Continuum **4**(6), 1776–1791 (2021).

16. H. M. Schoolman and L. M. Bernstein, "Computer use in diagnosis, prognosis, and therapy," Science **200**(4344), 926–931 (1978).

17. M. K. K. Niazi, A. V. Parwani, and M. N. Gurcan, "Digital pathology and artificial intelligence," The Lancet Oncology **20**(5), e253–e261 (2019).

18. R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," Radiology **278**(2), 563–577 (2016).

19. V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. W. L. Aerts, A. Dekker, and D. Fenstermacher, "Radiomics: the process and the challenges," Magn. Reson. Imaging **30**(9), 1234–1248 (2012).

20. H. J. W. L. Aerts, E. R. Velazquez, R. T. H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, and D. Rietveld, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," Nat. Commun. **5**(1), 4006 (2014).

21. P. Lambin, R. T. H. Leijenaar, T. M. Deist, J. Peerlings, E. E. C. De Jong, J. Van Timmeren, S. Sanduleanu, R. T. H. M. Larue, A. J. G. Even, and A. Jochems, "Radiomics: the bridge between medical imaging and personalized medicine," Nat. Rev. Clin. Oncol. **14**(12), 749–762 (2017).

22. H. Abdollahi, B. Mofid, I. Shiri, A. Razzaghdoust, A. Saadipoor, A. Mahdavi, H. M. Galandooz, and S. R. Mahdavi, "Machine learning-based radiomic models to predict intensity-modulated radiation therapy response, Gleason score and stage in prostate cancer," Radiol med **124**(6), 555–567 (2019).

23. S. K. Asrani, H. Devarbhavi, J. Eaton, and P. S. Kamath, "Burden of liver diseases in the world," J Hepatol **70**(1), 151–171 (2019).

24. P. Marcellin and B. K. Kutala, "Liver diseases: a major, neglected global public health problem requiring urgent actions and large-scale screening," Liver Int **38**(Suppl 1), 2–6 (2018).

25. P. Bedossa and F. Carrat, "Liver biopsy: the best, not the gold standard," J Hepatol **50**(1), 1–3 (2009).

26. P. J. Scheuer, "Classification of chronic viral hepatitis: a need for reassessment," J. Hepatol. **13**(3), 372–374 (1991).

27. P. Bedossa and T. Poynard, "An algorithm for the grading of activity in chronic hepatitis C," Hepatology **24**(2), 289–293 (1996).

28. K. Ishak, "Histological grading and staging of chronic hepatitis," J Hepatol **22**(6), 696–699 (1995).

29. K. P. Batts and J. Ludwig, "Chronic hepatitis: an update on terminology and reporting," The American Journal of Surgical Pathology **19**(12), 1409–1417 (1995).

30. T. Huang, R. Meng, J. Qi, Y. Liu, X. Wang, Y. Chen, R. Liao, and H. Ma, "Fast Mueller matrix microscope based on dual DoFP polarimeters," Opt. Lett. **46**(7), 1676–1679 (2021).

31. N. D. Theise, "Liver biopsy assessment in chronic viral hepatitis: a personal, practical approach," Mod. Pathol. **20**(S1), S3–S14 (2007).

32. P. Li, Y. Dong, J. Wan, H. He, T. Aziz, and H. Ma, "Polaromics: deriving polarization parameters from a Mueller matrix for quantitative characterization of biomedical specimen," J. Phys. D: Appl. Phys. **55**(3), 034002 (2022).

33. S.-Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on polar decomposition," J. Opt. Soc. Am. A **13**(5), 1106–1113 (1996).

34. N. Ghosh, M. F. G. Wood, and I. A. Vitkin, "Mueller matrix decomposition for extraction of individual polarization parameters from complex turbid media exhibiting multiple scattering, optical activity, and linear birefringence," J. Biomed. Opt. **13**(4), 044036 (2008).
35. P. Li, D. Lv, H. He, and H. Ma, "Separating azimuthal orientation dependence in polarization measurements of anisotropic media," Opt. Express **26**(4), 3791–3800 (2018).
36. A. Zwanenburg, S. Leger, M. Vallières, and S. Löck, "Image biomarker standardisation initiative," arXiv preprint arXiv:1612.07003 (2016).
37. F. J. Ferri, P. Pudil, M. Hatef, and J. Kittler, "Comparative study of techniques for large-scale feature selection," in *Machine Intelligence and Pattern Recognition* (Elsevier, 1994), pp. 403–413.
38. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, and V. Dubourg, "Scikit-learn: machine learning in Python," The Journal of Machine Learning Research **12**, 2825–2830 (2011).
39. A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems* (O'Reilly Media, 2019).